Empirical evidence for unique hues?

Bosten JM[1,2,*] snd Boehm AE[1,3]

[1]Department of Psychology, UC San Diego, McGill Hall, La Jolla, CA 92093
[2]Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge, CB2 3EB, United Kingdom
[3]Vision Science Group, UC Berkeley, 380 Minor Hall, Berkeley, CA 94720-2020
*Corresponding author: jmb97@cam.ac.uk

Received Month X, XXXX; revised Month X, XXXX; accepted Month X, XXXX; posted Month X, XXXX (Doc. ID XXXXX); published Month X, XXXX

Red, green, blue, yellow, and white have been distinguished from other hues as unique. We present results from two experiments that undermine existing behavioral evidence to separate the unique hues from other colors. In Experiment 1 we used hue scaling, which has often been used to support the existence of unique hues, but has never been attempted with a set of non-unique primaries. Subjects were assigned to one of two experimental conditions. In the "unique" condition, they rated the proportions of red, yellow, blue and green that they perceived in each of a series of test stimuli. In the "intermediate" condition, they rated the proportions of teal, purple, orange and lime. We find, surprisingly, that results from the two conditions are largely equivalent. In Experiment 2, we investigated the effect of instruction on subjects' settings of unique hues. We find that altering the color terms given in the instructions to include intermediate hues leads to significant shifts in the hue that subjects identify as unique. The results of both experiments question subjects' ability to identify certain hues as unique.

OCIS codes: (330.1690) Color; (330.1720) Color vision; (330.5510) Psychophysics; (330.5020) Perception psychology
http://dx.doi/org/10.1364/AO.99.099999

1. Introduction

Five hues: red, green, blue, yellow and white, have been given special status by color scientists as "unique." The unique hues are said to be phenomenologically pure, while all other hues are said to be phenomenologically mixed, containing various proportions of the five unique hues.

Unique hues have a long history in psychology. It was the belief that unique hues must have special status in color representation that led Hering [1] to propose red, green, blue and yellow as the poles of his opponent mechanisms. Following the tradition established by Hering, and later by Jameson and Hurvich [2], when color opponent cells were discovered in the primate lateral geniculate nucleus, they were described as being selective for red and green, and blue and yellow [3]. Though it is now clear that the chromatic tuning of color-opponent mechanisms revealed psychophysically does not map on to the Hering primaries [4], and that the color-sensitive cells in the lateral geniculate nucleus encode intermediate color directions [5-6], the conviction that unique hues must have prominence in neural color signals has led to persistence of the idea that color-opponent cells encode the Hering primaries, and to a search for higher-order color mechanisms that are selective for red, green, blue and yellow [7–10].

A. Hue scaling

Despite broad consensus among color scientists that red, green, blue and yellow should be considered unique, solid behavioral evidence to support this view has been hard to come by. One source of evidence that is often cited to justify the special status of unique hues [11-13], is hue scaling. Hue scaling as a method is based on sensory scaling of multiple stimulus attributes introduced by Jameson and Hurvich [14]. The subject is presented with colored stimuli and asked to consider how much of various given "primary" colors the stimulus contains. The subject assigns a numerical rating to each primary that reflects his judgment of how much of that primary is contained in his perception of the stimulus that is presented to him on a particular trial. The idea is that if a hue is unique, it will contain one primary, but not any other.

Results from hue scaling experiments typically show that there are four points around the hue circle where subjects report seeing only one primary. For example, at "unique" green, subjects report seeing green, but not blue, yellow or red. However, since the primaries to which subjects are asked to assign ratings are typically themselves the unique hues, it is perhaps unsurprising that ratings peak for red, green, blue and yellow in regions where neighboring hues approach the zero level. Red, yellow, green and blue are set apart roughly 90° around the hue circle, so it may be simply the perceptual distance between them that results in one not being seen in another.

Sternheim and Boynton [15] did hue scaling for monochromatic lights between 530 and 620 nm. They introduced orange as an available primary in some of their conditions, and concluded that they had evidence to separate yellow, green and red, but not orange, as unique. In four conditions subjects were given i) red, green and blue, ii) red, yellow and green, iii) red, orange and green, and iv) red, orange, yellow and green, to use as primaries. When yellow was unavailable as a primary, there were certain wavelengths that Sternheim and Boynton's subjects were unable to adequately describe by mixtures of red, green and blue. In contrast, when orange was unavailable, light of wavelengths in the "orange" part of the spectrum could be adequately described as mixtures of red and yellow.

We believe that Sternheim and Boynton's comparison between conditions was not entirely balanced. In conditions where orange was unavailable as a primary, two closely neighboring hues, red and yellow, were available. But in conditions where yellow was missing, at least one of the two alternative primaries provided was further from yellow on the hue circle than either red or yellow from orange. It could be simply be the spacing of the primaries in color similarity, rather than their uniqueness, that allows adequate description of other hues.

Neither Sternheim and Boynton, nor others who have since used hue scaling to investigate color appearance [16-20], have introduced a true control condition. In order to conclude that hue scaling results demonstrate the uniqueness of unique hues, it is necessary to show that the same results cannot be obtained using only non-unique primaries roughly evenly distributed around the hue circle. In Experiment 1, we report the results of hue scaling under two different conditions: a standard "unique" condition, where the four unique hues are given as primaries, and a control "intermediate" condition where four intermediate hues (purple, orange, lime, and teal) are given.

B. Effect of instruction on settings of unique hues

In experiments where subjects are required to identify or provide settings of unique hues, other unique hues are invariably given in the instructions [21-26]. A subject may be asked to set a yellow that is neither reddish nor greenish [27], or for measurements of unique green, to decide whether a given color is too bluish or too yellowish [17].

Subjects are required to identify a unique hue by consideration of other (typically unique) hues specified in the instructions. The two other unique hues contained in standard instructions are situated roughly orthogonally around the hue circle from the unique hue to be identified. This suggests possible artifacts. Red may be identified as neither bluish nor yellowish simply because blue and yellow are each sufficiently distant from red. Alternatively, red may be identified as neither blue nor yellow because it is perceptually half way between blue and yellow, rather than because it is subjectively pure. Our Experiment 2 controls for these potential experimental artifacts by substituting non-unique color terms into the instructions. Instead of being asked to identify a unique red that is neither bluish nor yellowish, subjects might be asked to identify a unique red that is neither bluish nor orangish. We predict that if subjects identify unique red as a single subjectively pure color, the change in instruction should have no effect on their "unique" hue settings.

2. Experiment 1: Hue scaling

A. Predictions

If unique hues are subjectively pure and other hues are subjectively mixed, what should the results of hue scaling be for our unique and intermediate conditions? Predictions are shown in Figure 1. For the unique primaries we expect functions like those that have been obtained by many other researchers. The function for each primary should peak at the hue angle of its associated unique hue, where the functions for the three other primaries should be at zero. At the positions of intermediate hues two primaries should be reported. The prediction for the intermediate primaries is very different. Functions for each intermediate primary should peak at the hue angle of that primary. But at the positions of the (subjectively unitary) unique hues all functions should be at zero.

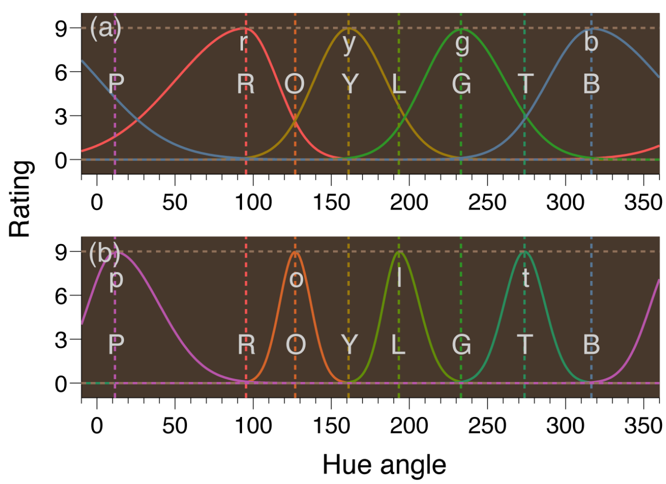

Figure 1. Predictions for hue scaling with unique primaries (panel (a)), and with intermediate primaries (panel (b)). Along the x-axis we plot hue angle around the hue circle, and on the y-axis the predicted numerical ratings from 0-9. The dashed vertical lines indicate the unique and binary hue primaries, labeled in uppercase: Purple (P), red (R), orange (O), yellow (Y), lime (L), green (G), teal (T) and blue (B). Predicted functions of ratings for each primary as a function of the hue angle of the test stimulus are indicated by the solid curves, labeled in lowercase. This figure is available in color online.

B. Stimuli

A representation of a particular trial is given in Figure 2(a). The test stimulus was a disk of 4° diameter presented in the upper half of the screen. The chromaticity of the test stimulus was randomly selected on each trial from 36 chromaticity coordinates distributed in a circle in MacLeod-Boynton [28] chromaticity space around a point metameric with illuminant D65 (Figure 2(b)). In the results section, the chromaticities of the 36 stimuli will be given by "hue angles". We defined hue angle in MacLeod-Boynton chromaticity space clockwise, with zero at the vertical (S/(L+M)) axis (Figure 2(b)). Conversion from hue angle to MacLeod-Boynton chromaticity coordinates can be achieved by the following formulae, where θ is the hue angle:

$$S / (L+M) = 0.0167 + 0.013\cos\theta$$
$$L / (L+M) = 0.6552 + 0.0364\sin\theta$$

The background on which the stimuli were presented was metameric with D65. All test stimuli were isoluminant, with a luminance of 28 cd m^{-2}, and the luminance of the background was 14 cd m^{-2}. In the lower part of the screen disks of 2° diameter were presented, containing four primaries (Figure 2(b)) which were either the unique hues or the intermediate hues, depending on the condition. The luminance of the primaries was 28 cd m^{-2}. Above each primary was a box in which the subject entered a number corresponding to his perception of how much of that primary was contained in the test stimulus. The boxes were delimited with a black line, one pixel wide. The subject used the arrow keys to move between boxes, and the currently selected box was indicated by replacing the black border with a white border, metameric with D65 and with a luminance of 28 cd m^{-2}. Subjects entered a rating into each box using a numerical keypad, and the number was shown in the box in white, also metameric with D65 and with a luminance of 28 cd m^{-2}.

C. Selection of primaries

The primaries used in the hue scaling experiment were based on selections of unique and intermediate hues made by 58 subjects [29]. Subjects were presented with annuli of selectable segments containing a progression of hues on a CRT monitor covered with a touch-sensitive screen. The segments containing the hues were isoluminant, with a luminance of 28 cd m^{-2}. The background was metameric with D65 and had a luminance of 17 cd m^{-2}. On each trial a circle of 90 segments containing 90 discrete hues was presented (Figure 2(c)). The circle was rotated randomly from trial to trial. According to the block, subjects were asked to choose, for example, “a red that is neither too orange nor too purple” or “an orange that is neither too red nor too yellow” [30]. The subject selected the segment he thought best matched the instruction, and a small achromatic disk appeared beside that segment. He confirmed his selection by tapping a check symbol presented in the lower left part of the screen.

There were 16 blocks, each of 15 trials. In each block one of the four unique hues (red, green, blue or yellow), or one of the four intermediate hues (orange, purple, teal or lime) was tested. In the first eight blocks, all eight hues were tested in a random order, and they were tested again in a different random order in the second eight blocks. 58 subjects took part in the experiment. All had normal color vision assessed using the Ishihara plates presented under MacBeth Illuminant C.

To gather subjects’ responses, we used a Keytec Magic Touch ProE-X touch screen (model number ET2032C) attached to a CRT monitor. Stimuli were presented on a Diamond Pro 2070SB CRT monitor (Mitsubishi, Tokyo, Japan) calibrated using a UDT photometer (United Detector Technology, Hawthorne, CA), and a SpectraScan PR650 spectroradiometer (Photo Research Inc., Chatsworth, CA). Experiments were run in Matlab (The MathWorks, Natick, MA), using a VSG2/4 graphics card (Cambridge Research systems, Rochester, UK).

The primaries used in the hue scaling experiment were based on the mean hue settings for the 58 subjects. These settings are shown in Figure 2(d) with 95% confidence intervals.

D. Procedure for hue scaling

Before beginning the experiment, subjects were told that their task was to make subjective decisions about the color of the test disk. Written instructions for the intermediate condition were as follows: “You will be shown a series of colors. For each color you must decide how much of four other colors your subjective experience of that color contains. You must decide how much orange, lime, purple and teal your experience of the color you are presented with contains. You must rate the quantity of each color in your experience by a number from 0-9. Enter your ratings into the boxes provided. You can navigate between boxes with the arrow keys. Accept all your ratings for a trial by pressing Return.” Equivalent instructions were given for the unique condition, except that the list of color terms was “red, yellow, blue and green”.

Subjects were given additional instructions orally. They were reminded that the task was subjective, that there was no right or wrong answer and that they must respond in whatever way they felt was most appropriate. They were instructed that the ratings they assigned did not have to add up to any particular number

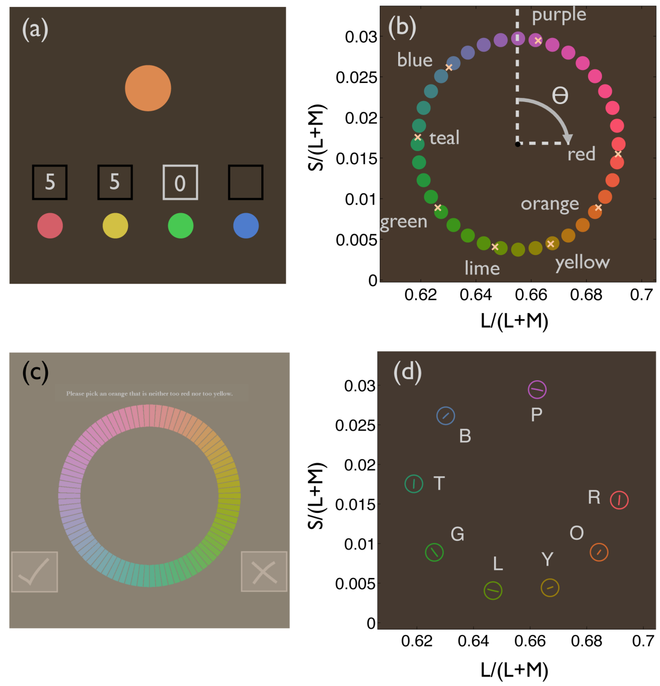

Figure 2. Stimuli used in the hue scaling experiment. Panel (a) represents the stimulus presented on one trial. The upper disk is the test stimulus, and the four lower disks are the primaries (left to right: Red, yellow, green and blue). The subject assigned a numerical rating to each primary according to how much of that primary he perceived in the test stimulus. The box currently selected was indicated with a white border. The stimulus represented here is for the unique condition: the stimulus for the intermediate condition was equivalent, except that the four primaries were teal, lime, orange and purple. Panel (b) shows the chromaticities of the stimuli in MacLeod-Boynton chromaticity space. The test chromaticities are shown by the disks, and the primaries by the gray crosses, labeled. The black dot in the center of the Figure indicates the chromaticity of D65. Hue angle (θ), used for the results, is indicated in the panel. Panel (c) represents the stimulus used for measuring unique and intermediate hues. Each of 90 segments contained a hue from a range linearly spaced around a circle centered at the coordinates of D65 in MacLeod-Boynton chromaticity space. Instructions were presented in the upper part of the screen – for the particular trial represented in the figure the instructions read “pick an orange that is neither too red nor too yellow”. The subject used a stylus to select a segment, and an achromatic disk appeared beside the selected segment. The selection was confirmed by tapping the check sign. Panel (d) shows mean selections of unique and intermediate hues made by 58 subjects. Data points for each primary are labeled: Purple (P), red (R), orange (O), yellow (Y), lime (L), green (G), teal (T) and blue (B). Error bars (lines inside data points) indicate 95% confidence intervals on the mean. The mean selections were used as the primaries in the hue scaling experiment. This figure is available in color online.

and that any possible response was acceptable, including giving all four primaries a score of zero.

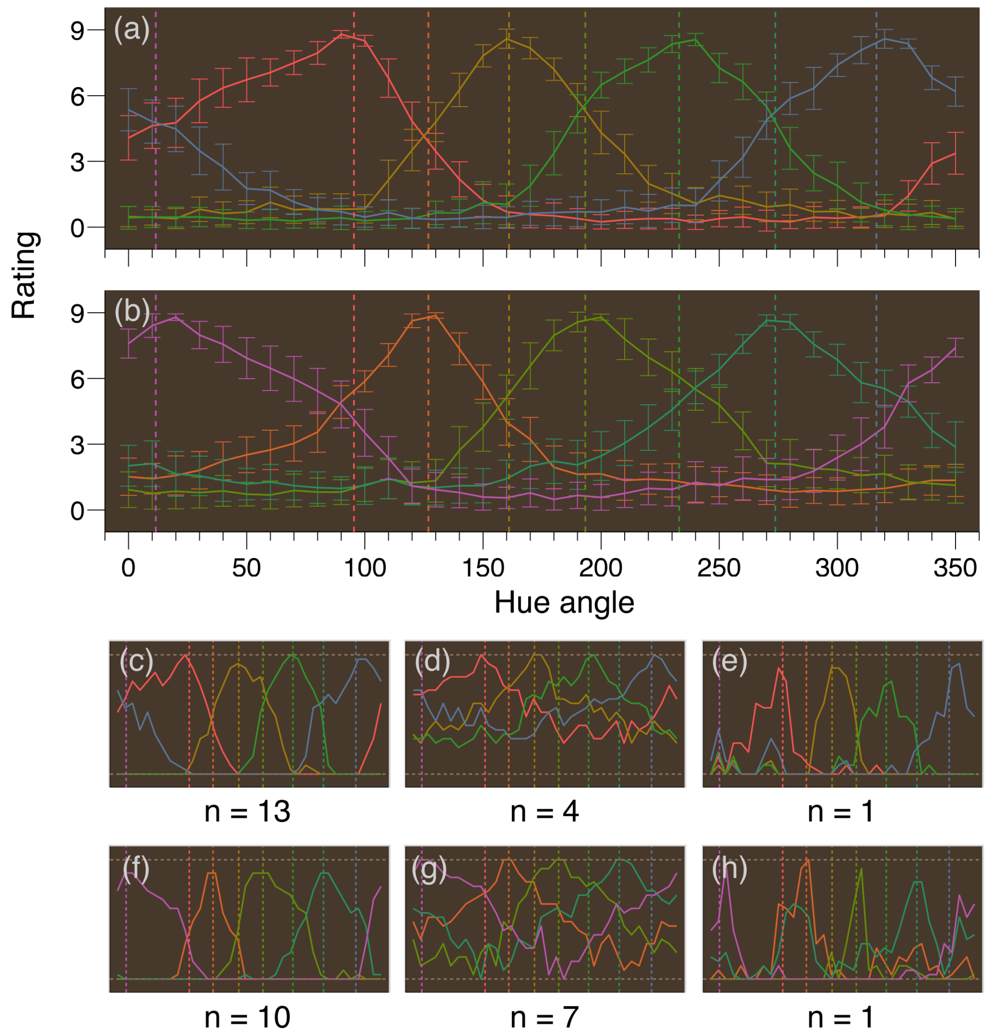

Figure 3. Results of the hue scaling experiment. In each panel, the hue angles (in degrees) of the test stimuli shown in Figure 2(b) are plotted along the x-axes and ratings are plotted up the y-axes. Each function shows ratings for how much of a particular primary subjects judged to be present in the test stimuli. Each panel shows four functions describing the results for the four primaries provided in that condition, colored according to the primary. Dashed vertical lines indicate the hue angles of the eight primaries used across the two conditions, also colored according to the primary. Panel (a) shows group mean results for the unique condition (n = 18), and panel (b) shows group mean results from the intermediate condition (n = 18). Error bars in panels (a) and (b) are 95% confidence intervals for each mean rating. Panels (c)-(h) show examples of results from individual subjects, with the n given in each case indicating the number of subjects who produced a result similar to the one shown. Panels (c), (d), and (e) are from the unique condition, and panels (f), (g) and (h) are from the intermediate condition.

On each trial a subject entered his ratings into the four boxes (Figure 2(a)). He was able to change his responses by overwriting them. When he was satisfied with all four ratings, he pressed the Return key to move on to the next trial. The 36 test chromaticities were presented in a random order three times, so that all 36 test stimuli were presented once in a random order, and then twice more in different random orders. A subject's rating for each test stimulus was based on the mean of the three trials. The experiment took most subjects between 10 and 30 minutes to complete.

E. Subjects

36 subjects took part in the experiment. All had normal color vision assessed using the Ishihara plates. Subjects were assigned randomly either to the "unique" or to the "intermediate" condition so that 18 subjects participated in each condition. The subjects were University of California, San Diego undergraduate students who were required to take part in experiments for course credit. They studied a range of subjects, including statistics, psychology, biology and linguistics. All subjects were naïve to the purposes of the hue scaling experiment and had not taken part in the experiment to select the chromaticities of the primaries.

F. Equipment

Stimuli were presented on a Mitsubishi Diamond Pro 2070SB CRT monitor. The gamma functions were linearized using measurements made with a UDT photometer calibrated for luminance using a Spectrascan PR560 spectroradiometer. The color calibration that allowed for conversion from MacLeod-Boynton chromaticity coordinates to RGB values was achieved by measuring the spectra of the three phosphors of the CRT monitor using the Spectrascan PR650.

G. Results

Results averaged across subjects are presented in Figure 3(a)-(b). The figure shows the mean ratings for how much of each primary was judged to be contained in each test stimulus. Results from the unique condition are presented in panel (a) and results from the intermediate condition in panel (b). In both panels, the hue angles of the eight primaries used in the two conditions are indicated by the vertical dashed lines.

Results of the unique condition are similar to those obtained by Sternheim and Boynton [15]. As Sternheim and Boynton found, peak ratings are at the hue angles of the unique hues. At the angle of each unique hue, ratings for the three other unique hues are very low. At the angles of intermediate primaries, two unique primaries are given medium ratings.

Suprisingly, the results of the intermediate condition are very similar to those of the unique condition. Peak ratings for each intermediate primary are found at the hue angle of that primary. At the angle of each intermediate primary, ratings for the three other intermediate primaries are low. At the angles of the unique primaries, two intermediate primaries are given medium ratings.

There are other notable features of the data. In each condition, the intersection of two descending functions occurs very near the hue angle of a primary not used in that condition. For example, in the unique condition, the functions for the red and yellow primaries cross at about 125°, which is very close to the hue angle of the orange primary that was not provided in the unique condition. In the intermediate condition, the functions for the lime and orange primaries cross at about 160°, which is close to the hue angle for the yellow primary that was not provided in the intermediate condition.

The most striking feature of the results is the similarity between the two conditions. Differences between conditions, by comparison, are relatively minor. However, the "baseline" ratings (ratings for primaries at test stimulus positions 180° removed from the position where ratings peak) are a little higher in the intermediate condition than in the unique condition. This small difference between conditions may be driven by individual differences. A minority of subjects performed unusually on this task, and it may be that in the two conditions there were different numbers of subjects who followed minority strategies.

H. Individual differences

Results from six individual subjects are shown in Figure 3(c)-(h). Results of the unique condition are presented in panels (c), (d) and (e), and results of the intermediate condition are presented in panels (f), (g) and (h). In each panel, n indicates the number of subjects who produced results similar to the example shown. Most subjects assign high ratings to one or two primaries for each test stimulus, and zero or near zero ratings to the other primaries (panels (c) and (f)). Other subjects, to varying degrees, assign non-zero ratings to more than two primaries for a given test stimulus (panels (d) and (g)). Two subjects (panels (e) and (h)) gave results that resembled the predictions for intermediate hues (Figure 1(b)), but one of these was given unique primaries.

The average variability in a rating is greater for the intermediate than for the unique condition: The mean standard deviation of ratings is 1.59 for the intermediate condition and 1.29 for the unique condition ($T = 5.9$, $p < 0.01$). However, rating variability only differs between conditions for low mean ratings (Figure 4). At medium and high mean ratings the distributions of variability in ratings overlap. This shows that the greater variability in the intermediate condition is only near the "baseline" level of rating. This greater variation in baseline ratings indicates that the slightly higher baseline ratings in the intermediate condition are caused by individual differences. Figures 3(c)-(h) bear this out: More subjects assign non-zero ratings to more than two primaries for a given test stimulus in the intermediate condition than in the unique condition.

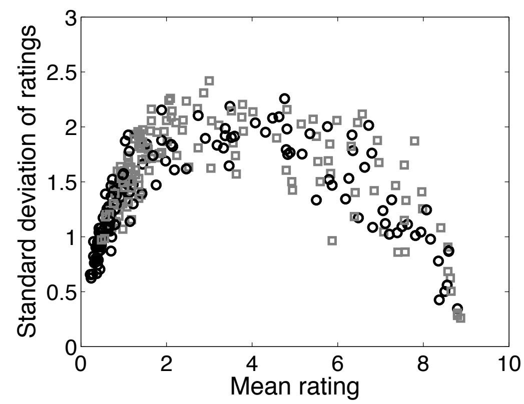

Figure 4. Distributions of rating variability with mean rating. Standard deviations of ratings for the intermediate condition are shown by the gray squares and for the unique condition by the black circles. The distributions of variability in ratings largely overlap. However, low ratings show lower variability in the unique condition than in the intermediate condition.

The small difference in results between the two conditions appears to be driven by individual differences in task strategy. There are clear examples of three different strategies in both conditions, (figures 3(c)-(h)), so we conclude that providing intermediate rather than unique primaries for hue scaling can lead to equivalent results.

3. Experiment 2: Effect of instruction on settings of unique hues

A. Stimuli

The stimulus was a disk of approximately 2° presented in the center of the display. The luminance of the disk was 28 cd m^{-2}. Its chromaticity varied according to a staircase procedure tracking the subject's responses. The starting positions of the staircases were at hue angles of 0° ($L/(L+M) = 0.655$, $S/(L+M) = 0.0297$), and 270° ($L/(L+M) = 0.619$, $S/(L+M) = 0.0167$) when unique blue was measured; and at 0° ($L/(L+M) = 0.655$, $S/(L+M) = 0.0297$) and 140° ($L/(L+M) = 0.679$, $S/(L+M) = 0.00675$) when unique red was measured. The background was metameric with D65 and had a luminance of 21 cd m^{-2}. Stimuli were presented on the same calibrated Mitsubishi Diamond Pro CRT monitor as in Experiment 1.

B. Procedure

Subjects were assigned randomly to one of three conditions for unique blue and to one of three conditions for unique red. Subjects were instructed that their task was to make subjective decisions about the color of the test disk. For unique blue they were then instructed (depending on the condition to which they were assigned) as follows: "In this block you are looking for a pure color that is"

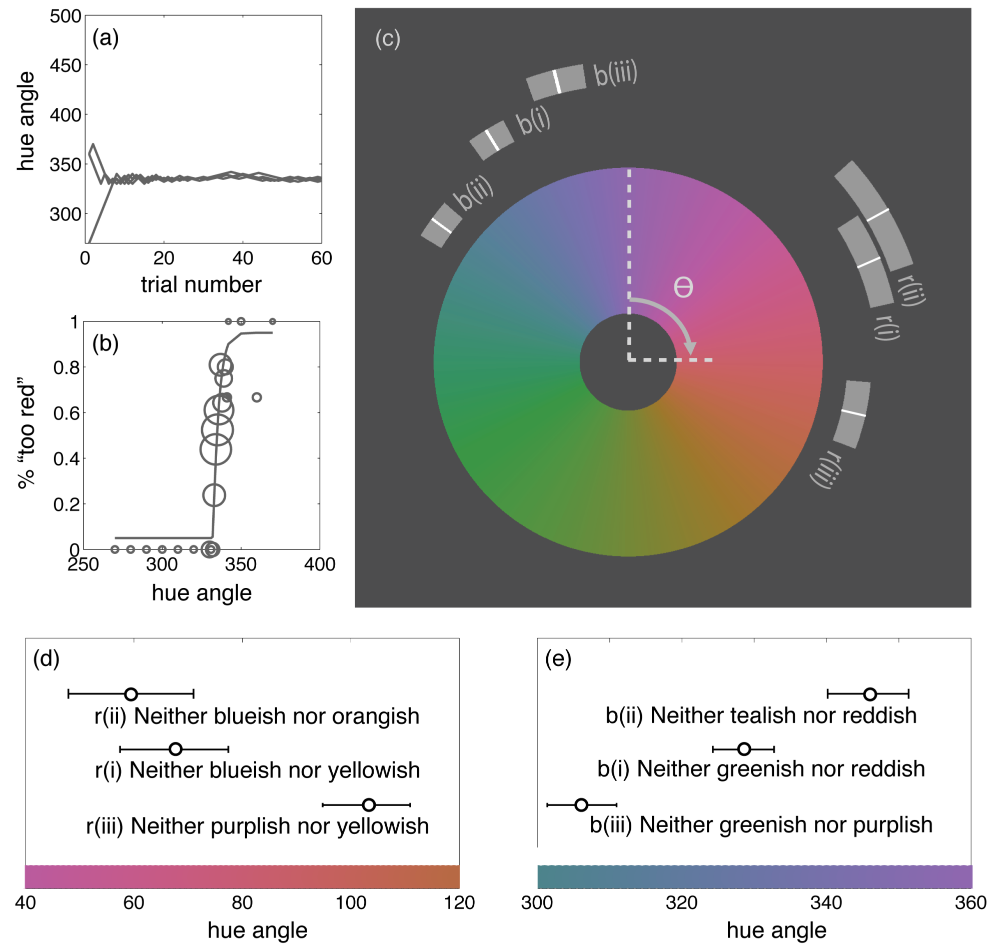

Figure 5. Results of Experiment 2. (a) Examples of staircases for one subject in condition b(ii): "Neither tealish nor reddish". (b) Cumulative Weibull psychometric function fit to the data shown in (a). The area of each data point is proportional to the number of times the particular stimulus was presented. (c) Group results. White bars indicate group means for each condition, and light gray areas indicate 95% confidence intervals. In each case the 'standard' version of the question (condition i) produces intermediate results, with alterations in the color words given in the instructions predictably shifting mean settings of unique hues (conditions ii and iii). (d) Results for unique red shown as a function of hue angle. Error bars are 95% confidence intervals. (e) Results for unique blue shown as a function of hue angle. Error bars are 95% confidence intervals. This figure is available in color online.

b (i) "neither greenish nor reddish"
b (ii) "neither tealish nor reddish"
b (iii) "neither greenish nor purplish"

They were instructed to press one button on a keypad if, for example, the disk was "too green", and another button if it was, for example, "too purple". For unique red, subjects were given equivalent instructions, except the color words given in the three conditions were:

r (i) blue and yellow
r (ii) blue and orange
r(iii) purple and yellow

Both unique blue and unique red were measured for each subject in separate blocks in a random order.

In each block four randomly interleaved staircases converged on a subject's unique hue. The initial step size was a hue angle of 10°, and this reduced to 5° once a pair of staircases had crossed (see Figure 5(a)). Each staircase terminated after 60 trials. There were 240 trials per block.

C. Subjects

115 Subjects took part in the experiment. None had taken part in Experiment 1. All had normal color vision assessed using the Ishihara plates. All subjects were naïve to the purposes of the experiment. They were University of California, San Diego undergraduates who participated in exchange for course credit.

D. Results

Results of Experiment 2 are shown in Figure 5. Panel (a) shows examples of staircases for unique blue. The subject whose data are shown in the panel was given instructions b(ii), to identify "a blue that is neither tealish nor reddish." Panel (b) shows a psychometric function fit to the data shown in Panel (a). We fit cumulative Weibull psychometric functions using *modelfree v1.1* [31]. Unique hue settings were defined as the 50% point on the psychometric function, where subjects were equally likely to give each of the two alternative responses. Panel (c) shows group results plotted in the context of the full hue circle. White lines show group mean hue settings, and the light gray areas indicate 95% confidence intervals. Hue angles of the hue settings are shown in panel (d) for unique red and panel (e) for unique blue.

The data show that mean settings of unique hues are shifted by changes in the instructions that are given to subjects. There was a significant effect of instruction both for unique red ($F = 22.7$, $p = 5.6 \times 10^{-9}$) and for unique blue ($F = 64$, $p < 1 \times 10^{-16}$). In both cases hue settings for the standard instructions were intermediate between those for the other two conditions. Compared to settings resulting from the standard instructions "a blue that is neither greenish nor reddish," settings of "a blue that is neither greenish nor purplish" are shifted away from purple and settings of "a blue that is neither tealish nor reddish" are shifted away from teal. The pattern of results for unique red is analogous to that for unique blue.

4. Discussion

A. Experiment 1: Hue scaling

We have found that, surprisingly, when a set of intermediate hues are substituted for the unique hues as primaries in hue scaling, results remain broadly the same. Our results are clearly against predictions of what should happen given the prevailing view that unique hues can be distinguished from other hues as subjectively pure (compare Figures 1 and 3). Subjects report seeing intermediate primaries in "unique" hues in the same way as they report seeing unique primaries in intermediate hues. In the intermediate condition, subjects assigned ratings of about five to each of the purple and orange primaries when presented with a red test stimulus. Similarly, in the unique condition, subjects assigned ratings of about five to each of the red and yellow primaries when presented with an orange test stimulus. If red, green, blue, and yellow really can be distinguished as unique, we should expect that when they are presented as test stimuli in the intermediate condition they would be given near-zero ratings (Figure 1(b)). This is clearly not evident in the results.

Is it possible that individual differences in unique hues affect our results? Individuals vary reliably in their settings of unique hues [21,30,32], though *within-subject* variability in hue settings accounts for a substantial portion of the total variability observed [29]. We used group mean unique and intermediate hue selections as primaries for our hue scaling task. The predictions shown in Figure 1(a) assume that the primaries provided are unique, but this might not be true for particular observers who have unique hue settings different from the mean. If, for a particular observer, the primaries provided in the "unique" condition were not unique, we would expect the pattern of results to resemble more the prediction for intermediate primaries shown in Figure 1(b). Since results for the majority of observers in both conditions matched the predictions for unique primaries (figure 1(a)), individual variability in hue settings cannot undermine our conclusion that the typical pattern of hue scaling results reported in the literature can be achieved with non-unique primaries.

Sternheim and Boynton [15] assessed from their hue scaling results whether a given hue was unique by applying three criteria: (a) that the color category should be used reliably; (b) that the rating function for a primary should reach a maximum in a region where functions for neighboring primaries are at a minimum; and (c) that a particular unique hue should not be represented in ratings assigned to other primaries when the primary associated with that unique hue is not permitted. All the primaries we used in both unique and intermediate conditions satisfy Sternheim and Boynton's first two criteria (see Figure 3). Take the function for orange as an example: The category is used reliably as evidenced by small confidence intervals and consensus across subjects, and the function reaches a maximum in a region where neighboring hues (violet and lime) are at a minimum. No primaries from either condition fulfill Sternheim and Boynton's third criterion. The color primaries "red", "green", "blue" and "yellow" were not permitted in the intermediate condition. Yet when red, green, blue or yellow test stimuli were presented, subjects were able to make responses using the intermediate primaries they were provided with. "Unique" yellow, for example, was described by assigning ratings of about 5 to each of the permitted primaries of orange and lime.

A minority of color scientists have questioned the assumed special status of unique hues. Jameson [32] has done so by considering the variable responses of participants in the World Color Survey [33]. A paper by Saunders and van Brakel [34], questioning the linguistic, neurophysiological and psychophysical evidence for unique hues triggered many energetic responses. They were accused of "throwing the baby out with the bath water" [35] and "selective vision" [36]. They made a valuable point, however, that solid behavioral evidence for the existence of unique hues is lacking. Hue scaling results were put forward by several respondents as evidence supporting the distinction between unique and intermediate hues [13, 35–37]. Results from the present study undermine this evidence. Broakes [38], also in response to Saunders and van Brakel, challenged color scientists to perform experiments using lime, purple, orange and teal, and suggested that "The fears of Saunders and van Brakel will be justified if people can prove to do as well with lime, purple, orange and teal as we do with red, yellow, blue and green". For hue scaling, surprisingly, this turns out to be the case.

What can explain subjects' behavior in our hue scaling experiment? One possibility is that the assumption that there is consensus for the subjective uniqueness of red, green, blue and yellow is misguided. But another possibility, which is impossible to rule out in a subjective task of this nature, is that subjects ignored or misinterpreted the instructions and instead performed a color similarity task, rating the similarity of the primary and the test hue. But if this could explain the results of the present study, it could also explain the form of all other results from hue scaling where unique hues are given as primaries. Conclusions from hue scaling, one of the main lines of evidence that has been advanced in support of the special status of unique hues, are unsound.

B. Experiment 2: Effect of instructions on settings of unique hues

We have found that substituting intermediate color terms into the instructions alters mean settings of unique hues. If subjects are asked for a unique red that is neither purplish nor yellowish, they identify a much more orange shade of red than if they are asked for a unique red that is neither bluish nor yellowish. Our sample of naïve subjects, despite in most cases producing clean psychometric functions (e.g. Figure 5(b)), do not reliably identify a particular hue as subjectively pure. Our results allow for the possibility that unique hues do not exist, and that settings of unique hues might simply reflect a perceptual intermediate between the alternatives given in the instructions. However, we cannot exclude the possibility that our subjects were simply biased by the contents of the instructions. What we *can* conclude is that deciding whether a particular hue is neither "too red" nor "too green" to be unique is not a simple enough or salient enough task that subjects can escape bias. We can also conclude that the fact that subjects are able to identify a particular blue as neither greenish nor reddish is not evidence that the blue is unique.

C. General Discussion

Our results undermine existing behavioral evidence for the special status of unique hues. But our experiments, and most others that have attempted to establish the uniqueness of unique hues empirically, employ Brindley class B observations [39], where subjects are required to describe the quality of their sensations. One recent and particularly interesting example of a class B method of investigating unique hues is *partial hue matching* by Logvinenko and Beattie [40], where subjects are required to identify the set of hues that shares a component color quality with a test hue. *Chromaticity classes* (the largest set of stimuli which all partially match) are extracted from matrices of partial hue matches, and for about 50% of observers four "unique" hues that belong to only one chromaticity class emerge. However, there are several minority phenotypes, including subjects who seem to have more than four unique hues, or who have no unique hues. The difficulty of interpreting the possible alternative strategies of subjects with atypical results has led Logvinenko [41] to suggest that "this problem is far too involved to be left to inexperienced observers." However, it seems equally likely that observers who are experienced in making color introspections might give invalid or unrepresentative results. Even those who are not versed in color science may adopt an introspective strategy of basing subjective judgments on the "primary" colors that they are culturally familiar with.

Given the limitations of methods involving introspection, could any psychophysical experiment either prove or disprove the existence of unique hues? What about performance measures, or class A observations?

Subjects' reliability at setting unique hues is not easily categorized either as a class B or as a class A observation. In the *method of average error* (which constitutes a class A observation), subjects are required to make a match many times, and the variance of matches is taken as a measure of performance. Measurement of the reliability of settings of unique hues uses the method of average error, but the standard to which subjects are making a 'match' is internal rather than external, so subjects are required to judge the quality of their sensations. In either case, it might be thought reasonable to assume that subjects can set unique hues more reliably than they can set intermediate hues. However, Malkoc et al. [30] found that test-retest reliability is as high for binary (intermediate) hues as it is for unique hues. Bosten and Lawrance-Owen [29] have independently replicated Malkoc et al.'s findings and found that intra-subject subject variability (as well as inter-subject variability) for identifying unique hues is no different to that for identifying intermediate hues.

Though it is perhaps telling that pure class A observations that distinguish the unique hues have, until recently, been lacking (some studies have reported negatively, that unique hues are not distinguished from other hues using class A observations [29,42–46]), Danilova and Mollon [47] have found that the locus of unique blue and unique yellow corresponds to minima in thresholds for color discrimination. The correlation between the loci of unique hues, measured using subjective judgments, and a peak in discrimination performance is interesting. Such findings do, however, need to account for the wide range of settings of unique hues found across individuals and across studies [48]. A correlation between individual differences in the positions of color discrimination minima and individual differences in settings of unique hues would be provocative.

Danilova and Mollon's careful measurements, and other Brindley class A observations, might be able to establish the presence or absence of channels tuned to particular chromaticities, but they can no more establish the uniqueness of unique hues than the more widespread class B observations. Uniqueness is defined subjectively, and therefore cannot be proven. Perhaps it is time for color scientists to put unique hues aside, and know them simply as red, green, blue and yellow.

Acknowledgements

This work was supported by a Research Fellowship from Gonville and Caius College, and by grant EY011711 from the National Institutes of Health awarded to Donald MacLeod.

References

1. E. Hering, *Zur Lehre Vom Lichtsinne* (Carl Gerold's Sohn, 1878).
2. L. M. Hurvich and D. Jameson, "An opponent-process theory of color vision," Psychol. Rev. 64, 384–404 (1957).
3. R. De Valois, I. Abramov, and G. Jacobs, "Analysis of response patterns of LGN cells," J. Opt. Soc. Am. 56, 966–977 (1966).
4. J. Mollon and C. Cavonius, "The chromatic antagonisms of Opponent Process Theory are not the same as those revealed in studies of detection and discrimination," in *Colour Vision Deficiencies VIII*, G. Verriest, ed. (Junk, 1987).
5. K. A. Jameson and R. G. D'Andrade, "It's not really red, green, yellow, blue: an inquiry into perceptual color space," in *Color Categories in Thought and Language*, C. Hardin and L. Maffi, eds. (Cambridge University Press, 1997), pp. 295–319.
6. I. Abramov and J. Gordon, "Color appearance: on seeing red-or yellow, or green, or blue," Annu. Rev. Psychol. 3, (1994).
7. J. Mollon, "A neural basis for unique hues?," Curr. Biol. 19, R441–R442 (2009).
8. C. M. Stoughton and B. R. Conway, "Neural basis for unique hues.," Curr. Biol. 18, R698–9 (2008).
9. S. M. Wuerger and L. Parkes, "Unique hues: Perception and brain imaging," in *New Directions in Colour Studies*, C. Biggam, C. Hough, C. Kay, and D. Simmons, eds. (John Benjamin Publishing Company, 2011), pp. 445–455.
10. A. Valberg, "Unique hues: an old problem for a new generation.," Vision Res. 41, 1645–57 (2001).
11. D. R. Hilbert, "Basic tastes and unique hues," Behav. Brain Sci. 31, 82 (2008).
12. M. Danilova and J. Mollon, "Foveal color perception: Minimal thresholds at a boundary between perceptual categories," Vision Res. 62, 162–172 (2012).
13. D. Miller, "Over the rainbow: The classification of unique hues," Behav. Brain Sci. 20, 204–205 (1997).
14. D. Jameson and L. M. Hurvich, "Perceived color and its dependence on focal, surrounding, and preceding stimulus variables.," J. Opt. Soc. Am. 49, 890–898 (1959).
15. C. E. Sternheim and R. M. Boynton, "Uniqueness of perceived hues investigated with a continuous judgmental technique.," J. Exp. Psychol. 72, 770–6 (1966).

16. R. L. De Valois, K. K. De Valois, E. Switkes, and L. Mahon, "Hue scaling of isoluminant and cone-specific lights.," Vision Res. 37, 885–97 (1997).
17. V. J. Volbrecht, J. L. Nerger, L. S. Baker, A. R. Trujillo, and K. Youngpeter, "Unique hue loci differ with methodology," Ophthalmic Physiol. Opt. 30, 545–552 (2010).
18. I. Abramov and J. Gordon, "Seeing unique hues," J. Opt. Soc. Am. A 22, 2143–2153 (2005).
19. J. Gordon, I. Abramov, and H. Chan, "Describing color appearance: hue and saturation scaling," Percept. Psychophys. 56, 27–41 (1994).
20. B. Wooten and D. Miller, "The psychophysics of color," in *Color Categories in Thought and Language*, C. Hardin and L. Maffi, eds. (Cambridge University Press, 1997), pp. 59–88.
21. M. A. Webster, E. Miyahara, G. Malkoc, and V. E. Raker, "Variations in normal color vision. II. Unique hues.," J. Opt. Soc. Am. A. Opt. Image Sci. Vis. 17, 1545–55 (2000).
22. D. Hinks, L. M. Cárdenas, R. G. Kuehni, and R. Shamey, "Unique-hue stimulus selection using Munsell color chips.," J. Opt. Soc. Am. A. Opt. Image Sci. Vis. 24, 3371–8 (2007).
23. G. Jordan and J. D. Mollon, "Rayleigh matches and unique green.," Vision Res. 35, 613–20 (1995).
24. F. Dimmick, "The spectral location of psychologically unique yellow, green and blue," Am. J. Psychol. (1939).
25. B. E. Schefrin and J. S. Werner, "Loci of spectral unique hues throughout the life span.," J. Opt. Soc. Am. A. 7, 305–11 (1990).
26. J. M. Eichengreen, "Unique hue loci: induced shifts with complementary surrounds.," Vision Res. 16, 199–203 (1976).
27. D. H. Brainard, a Roorda, Y. Yamauchi, J. B. Calderone, a Metha, M. Neitz, J. Neitz, D. R. Williams, and G. H. Jacobs, "Functional consequences of the relative numbers of L and M cones.," J. Opt. Soc. Am. A. Opt. Image Sci. Vis. 17, 607–14 (2000).
28. D. I. MacLeod and R. M. Boynton, "Chromaticity diagram showing cone excitation by stimuli of equal luminance.," J. Opt. Soc. Am. 69, 1183–6 (1979).
29. J. Bosten and A. J. Lawrance-Owen, "No difference in variability of unique hue selections and binary hue selections," Submitt. to JOSA (n.d.).
30. G. Malkoc, P. Kay, and M. A. Webster, "Variations in normal color vision. IV. Binary hues and hue scaling.," J. Opt. Soc. Am. A. Opt. Image Sci. Vis. 22, 2154–68 (2005).
31. K. Zychaluk and D. H. Foster, "Model-free estimation of the psychometric function.," Atten. Percept. Psychophys. 71, 1414–25 (2009).
32. K. A. Jameson, "Where in the World Color Survey is the support for the Hering Primaries as the basis for Color Categorization?," in *Color Ontology and Color Science*, J. Cohen and M. Matthen, eds. (The MIT Press, 2010), pp. 179–202.
33. P. Kay and T. Regier, "Resolving the question of color naming universals.," Proc. Natl. Acad. Sci. U. S. A. 100, 9085–9 (2003).
34. B. A. Saunders and J. van Brakel, "Are there nontrivial constraints on colour categorization?," Behav. Brain Sci. 20, 167–179 (1997).
35. I. Abramov and J. Gordon, "Constraining color categories: The problem of the baby and the bath water," Behav. Brain Sci. 20, 179–180 (1997).
36. M. Bornstein, "Selective vision," Behav. Brain Sci. 20, 180–181 (1997).
37. A. Byrne and D. Hilbert, "Unique hues," Behav. Brain Sci. 20, 184–185 (1997).
38. J. Broakes, "Could we take lime, purple, orange, and teal as unique hues?," Behav. Brain Sci. 20, 183–184 (1997).
39. G. Brindley, *Physiology of the Retinal and Visual Pathway. Second Edition.* (The Camelot Press Ltd., 1970).
40. A. D. Logvinenko and L. L. Beattie, "Partial hue-matching," 11, 1–16 (2011).
41. A. D. Logvinenko, "A theory of unique hues and colour categories in the human colour vision," Color Res. Appl. 37, 109–116 (2012).
42. M. A. Webster and J. D. Mollon, "Changes in colour appearance following post-receptoral adaptation.," Nature 349, 235–8 (1991).
43. C. Witzel and K. R. Gegenfurtner, "Categorical sensitivity to color differences," 13, 1–33 (2013).
44. S. Burns, A. Elsner, J. Pokorny, and V. Smith, "The Abney effect: chromaticity coordinates of unique and other constant hues.," Vision Res. 24, 479–89 (1984).
45. G. Malkoc and F. A. A. Kingdom, "Dichoptic difference thresholds for chromatic stimuli.," Vision Res. 62, 75–83 (2012).
46. M. Ayama, T. Nakatsue, and P. K. Kaiser, "Constant hue loci of unique and binary balanced hues at 10, 100, and 1000 Td.," J. Opt. Soc. Am. A. 4, 1136–44 (1987).
47. M. V Danilova and J. D. Mollon, "Cardinal axes are not independent in color discrimination.," J. Opt. Soc. Am. A. Opt. Image Sci. Vis. 29, A157–64 (2012).
48. R. Kuehni, "Variability in unique hue selection: A surprising phenomenon," Color Res. Appl. 29, 158–162 (2004).